

# Crossing Empirical Trend Analysis (CETA) At Risk Levels In Hydro-Meteorological Time Series

Zekai Sen (✉ [zsen@medipol.edu.tr](mailto:zsen@medipol.edu.tr))  
Istanbul Medipol Universitesi

---

## Research Article

**Keywords:** Crossing, empirical, risk, trend, hydro-meteorology, monotonic, median, multi

**Posted Date:** April 2nd, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-372087/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

# Abstract

Trend identification procedures are employed to determine the systematic monotonic trend lines in a given hydro-meteorological time series records for depiction of time dependent changes in the form of increase or decrease. Different methodologies are proposed for such identifications, but most of them require restrictive assumptions such as the normal (Gaussian) probability distribution, serial independence and long sample sizes. In order to relieve especially the serial independence requirement pre-whitening and over-whitening procedures are suggested, but they cannot render a serially dependent series into completely independent structure. In this paper, a new trend methodology is proposed on the basis of crossing features along any given straight-line within the given time series and the one with the maximum crossing number is the searched trend component. This approach does not require any restrictive assumption. Contrary to the previous trend algorithms, the suggested crossing empirical trend analysis (CETA) yields not a single trend, but a set of trends at different levels within the variation range of hydro-meteorological time series records. In this paper for the sake of brevity only three levels are considered at 10%, 50% and 90% risk levels. The comparison of the CETA approach is presented with the classical and frequently used method of Mann-Kendall (MK) trend identification procedure based on the Sen's slope calculation. For small serial correlation coefficients and normal probability distribution (PDF) function cases CETA and classical technique yield almost the same trend line within +5% error band limits. The application of this methodology is presented for monthly and annual discharge records of Danube River and annual precipitation records from seven geographical regions of Turkey.

## 1. Introduction

As mentioned by Milly et al. (2008) the stationarity of hydro-meteorological time series is dead and non-stationarity features have started as the most attractive components in the forms of trend due to the global warming and its consequent climate change impacts. The design and operation of various hydro-meteorological activities on the basis of stationarity assumption implies that the future is the reflection of the past. Global warming and climate change impacts do not follow such a stationarity evolution and any calculation based on such an assumption is bound to lead to over- or under-design conclusions, and hence, water resources systems cannot serve their purposes adequately. In order to avoid such inadequate design consequences it is necessary to search for strong natural and anthropogenic systematic changes identifications such as trends and variabilities (extreme events) through both reliable data and effective methodologies.

Trend identification methodologies appear as in the form of parametric or non-parametric procedures each with a set of some restrictive assumptions. Among the non-parametric trend identification tests are Spearman Rho (SR) and Mann-Kendall (MK) (Mann, 1954; Kendall, 1975) methods. The MK test requests serially independent structure, but almost all hydro-meteorological data have significant serial autocorrelations (Khaili et al, 2016). In cases of serial autocorrelation, the MK trend identification test may lead to even no trend result (Ahmadi et al., 2018). The reasons for the non-parametric tests usage instead of parametric methods is explained in detail by Hirsch et al. (1991). Prior to the SR test application, the data is sorted in ascending chronological order and a rank number is assigned to each data point. Subsequently, depending on the ranks and sample number a test statistics is defined (Yue et al., 2002). Another alternative for trend identification is the Innovative Trend analysis (ITA) method as proposed by Şen (2012), which does not require restrictive assumptions.

Non-parametric methods are applicable to both the stationary and non-stationary time series (Khalili et al., 2016). As for the parametric method, the linear regression approach is the most frequently used for analyzing the relationship between hydro-meteorological variable and the time of occurrence. This method has a set of restrictive assumptions such as the normal (Gaussian) probability distribution function (PDF) and independence of the residuals, which are not quite satisfied by natural event records. As for the outliers, the parametric methods are very sensitive, which is not the case in the non-parametric methodologies. Among the trend analysis methodologies no assumption requirement methods are distribution-free about the underlying distribution of the data, but the independence assumption still remains as the major problem.

Hamed and Rao (1998) investigated the effect of autocorrelation in time series on trend detection by non-parametric tests. They proposed a methodology for eliminating the effect of autocorrelation from a data series and applied the proposed

method for identifying trends in precipitation and stream flow time series showing that after eliminating the autocorrelation effect the MK test provides an increase in the accuracy of trend detection, but still there are uncertainties.

The majority of hydro-meteoro-climate time series are non-normally distributed, and therefore, the use of distribution-free methods usages makes sense. Although distribution-free methods require no assumptions about the underlying distribution of the data, but the independence assumption remains as the major problem. On the other hand, parametric trend identification tests are based on normal PDF requirement and independence assumptions.

Extreme values trend testing is among the non-stationarity procedures that is very difficult to identify. There are different classical options for their separation and then statistical treatment. This paper proposes a new methodology for their analysis on the basis of risk level assessment. Among the classical extreme value procedures are two approaches such as the selection of the largest value in each year, where the some of the extremes are insufficiently accounted and peaks over a threshold extremes, which are referred to as the partial duration series, where the extreme values are supposedly considered as independent. Both extreme values selection has some subjective ingredients, but the crossing empirical trend analysis (CETA) methodology of this paper provides an objective procedure for threshold selection in the form of trends with slopes rather than a constant level as in the partial duration series construction or as annual extreme selections. This paper presents a trend identification methodology for extreme high and low time series values without any distinctive separation as in the classical extreme value analysis. In the literature the crossing trend analysis has been presented first time by Şen (2018), but as a single monotonic trend with Sen (1968) slope calculation. However, in this paper completely new slope calculations are presented for a multitude of levels.

The main purpose of this paper is to propose a new trend analysis approach based on the crossing number along a set of straight-lines. The best representative trend line is selected among the set as the one with maximum crossing number. The more the crossing points along a line, the less is the serial correlation coefficient. This methodology is referred to as the crossing trend analysis (CETA) approach, which can be applied at any level inside the hydro-meteorology variation domain. The implementation of CETA can be achieved for a set of risk levels, but for the sake of brief explanation, herein only three risk levels are considered, namely, median (50%) and two extreme levels (10% and 90%). The application of the CETA methodology is presented for annual and 12 monthly discharge records from Danube River runoff and annual rainfall records from seven geographical regions of Turkey.

## 2. Crossing Empirical Trend Analysis (Ceta) Methodology

In general, trend appears as a straight line fit along the time axis to the data set with slope different from zero. Hirsch (1991) has suggested that for monotonic trend analysis minimum five years of monthly data are necessary.

The basic concept in the CETA methodology is that there are a set of straight lines with different locations and slopes inside the variation domain of a hydro-meteoro-climatic time series records. The main problem is to identify the one that is the most representative for general tendency. As for the parametric trend search methods such as the linear regression line, the location is the centroid point, i.e., the arithmetic average of the records and the slope is related to the regression coefficient. In the MK trend test the same location appears but with median slope value as suggested by Sen (1968). The location is the arithmetic average in case of normal PDFs, where the arithmetic average, median and mode collapse on each other, but in skewed PDF cases, arithmetic averages may not be a convenient choice. As mentioned earlier, these methodologies are dependent on a set of assumptions among which are the most effective ones as the normal (Gaussian) PDF and theoretically zero serial correlation coefficient.

The basis of the CETA approach is for the identification of the representative trend straight-line by completely different methodology, where the crossing numbers of any straight line with the record trace is sought. Although there are numerous crossing number depending on the location of the straight line, but the one with the maximum crossing number corresponds to the best representative trend line. Such an empirical approach does not require any assumption. Figure 1 shows an

illustrative time series for the explanation of CETA methodology. There are two most extreme upper and lower trends, which are determined simply by connecting the maxima (minima) between the two halves of the given time series.

It is necessary to consider the following steps for the explanation of the CETA methodology for trend identification.

1) Consider the two halves of the given time series and identify in each half maximum and minimum data point values and their record times. In Figure 1 the four points are shown as  $(X_{U1}, Y_{U1})$ ,  $(X_{U2}, Y_{U2})$ ,  $(X_{L1}, Y_{L1})$  and  $(X_{L2}, Y_{L2})$ ,

2) Calculate the slopes of each extreme trend lines as  $S_U$  and  $S_L$  by consideration of the symbols in the previous step as,

$$S_U = \frac{Y_{U2} - Y_{U1}}{X_{U2} - X_{U1}} \quad (1)$$

and

$$S_L = \frac{Y_{L2} - Y_{L1}}{X_{L2} - X_{L1}} \quad (1)$$

3) Chose a pivot point on the vertical half for depicting a set of straight-lines with different slopes between  $S_L$  and  $S_U$ . It is possible to choose any point on this vertical as a pivot such as the arithmetic average, median or any value between the upper and lower extreme boundary trends. In the application part of this paper, median is chosen as the pivot point, because it corresponds to the 50% probability of any PDF, i.e. 50% risk level,

4) Among all the different straight-lines slopes, the one with the maximum crossing point is the valid trend line. Herein, the maximum crossing number implies the least serial correlation coefficient, because the more is the crossing number the less is the serial correlation coefficient (Şen, 2018).

5) The slope of the empirically identified monotonic straight-line is calculated as the tangent of the line, which provides a completely different slope calculation than the classical Sen (1968) slope.

### 3. Study Area And Data

As for the study area two regions are adapted, which are the Danube River catchment monthly and annual average discharge measurements at Orshova gauge station with records starting from 1840 until 2013. This river has the second largest catchment area (801,463 km<sup>2</sup>) in Europe and it is the world's most international river basin, flowing through 19 countries with confluence at the Blach Sea coast in Romania. The second set of data are annual precipitation records from seven different geographical regions of Turkey each with the major city (Istanbul/Florya, Izmir, Adana, Sanliurfa, Van, Konya and Samsun) as in Fig. 2.

### 4. Applications

Long time change detections in river flow records are practically important research issue, because rivers are the most important and easily accessible water resources. Danube river monthly and annual discharge records are examined for possible trend identifications in this paper. Although there are discharge records for each month, herein season representative months are considered as February, May, August and November in addition to the annual discharges. Prior to the CETA's application each month PDF is depicted as presented in Fig. 3. Table 1 indicates some specific features of each month and annual discharge records.

Each one of the graphs has different PDF with parameters that are not close to each other. There is no normal (Gaussian) PDF, which is the requirement of some classical trend analysis procedures. The theoretical PDFs in this figure are matched to empirical data and the maximum difference between the two remains below the Kolmogorov-Smirnov test. This test is used to decide whether a sample comes from a population with a specific PDF (Chakravart et al., 1967). In the same graphs, 50%

median level is read from the horizontal axes with the upper and lower risk levels corresponding to two extreme risk levels and the values are given in the last three columns of Table 1.

The serial independence test is sought through the autocorrelation coefficient, but since none of the available time series abide with the normal (Gaussian) PDF, the results can be considered as biased values, because the autocorrelation coefficient is valid only for Gaussian PDFs. In Table 1 first, second and third lag correlation coefficient values are shown, and one can assume that all of them have almost independent serial correlation. This point may be considered without any pre-whitening or over-whitening procedure application (Bayazit and Önöz, 2007; Şen, 2018).

Table 1  
Danube River discharge statistical parameters

Month	PDF Type	Statistical parameters			Correlation coefficient			Lower	Median	Upper
		Location	Scale	Shape	First	Second	Third	10%	50%	10%
January	Gamma	6.568	733.423	-	0.037	0.078	0.092	2593	4563	7252
February	Pearson	-0.147	1519.392	4396.757	0.008	0.123	0.083	2619	4575	7230
March	Pearson	-0.152	1711.059	5766.054	0.106	-0.044	-0.04	3016	4913	7252
April	Pearson	-0.182	1988.33	7014.232	-0.100	-0.038	-0.046	5111	7689	10611
May	Log-normal	8.915	0.255	-	-0.034	0.071	0.069	5224	7719	10521
June	Gamma	15.270	437.423	-	-0.026	0.142	0.094	5285	7403	10184
July	Pearson	-0.035	1292.277	4886.303	0.019	-0.07	0.042	3765	5313	7550
August	Log-normal	8.371	0.299	-	0.031	0.034	-0.124	2943	4321	6270
September	Log-normal	8.251	0.297	-	0.086	0.042	0.072	2586	3815	5555
October	Gamma	9.586	422.801	-	0.075	0.094	0.107	2450	3911	5660
November	Pearson	0.052	1374.533	3949.332	0.069	0.005	0.048	2827	4458	7133
December	Pearson	-0.069	1516.943	4370.007	0.0375	0.078	0.092	2898	4891	7521
ANNUAL	Log-normal	8.61	0.169	-	0.203	0.029	0.108	4419	5488	6771

The same time series are treated with CETA methodological trend identification procedure at three risk levels in Figure 4.

Similar CETA procedure calculations are affected for the annual precipitation time series records from different parts of Turkey and the risk graphs are shown with the PDFs in Fig. 5. The corresponding PDF types and their statistical parameters are given in Table 2. None of the stations have symmetric normal (Gaussian) PDF, but they are all skewed to a certain extent, which is against the normal PDF requirement of the MK trend identification test.

In this figure the 50% median, 10% upper and lower risk levels obtained from the theoretical PDFs are given in the last three columns in Table 2. As for the first order serial correlation coefficients all passed from Kolmogorov-Smirnov test except three locations, namely, Izmir, Konya and Samsun meteorology stations.

Table 1  
Turkish precipitation records statistical parameters

Location	PDF type	Statistical parameters			Correlation coefficient			Lower	Median	Upper
		Location	Shape	Scale	First	Second	Third	10%	50%	10%
Florya	Pearson	-0.105	98.979	597.994	0.07	-0.191	-0.052	511.7	633.6	790.1
Izmir	Log-normal	6.509	0.263		0.102	-0.085	-0.02	479.2	671.5	928.5
Konya	Pearson	-0.157	69.424	293.762	0.285	0.0148	-0.003	231.8	318.5	421.6
Şanlıurfa	Gamma	13.357	34.383		-0.075	0.026	-0.112	307.4	447.9	619.2
Van	Pearson	-0.23	66.199	359.496	0.07	0.032	-0.017	298.6	382.8	478.7
Adana	Gamma	13.919	46.759		-0.013	0.062	-0.101	439.6	635.3	875.4
Samsun	Pearson	-0.13	103.076	659.968	0.104	-0.01	0.014	569.1	696.8	857.2

In Figure 6, CETA methodology median, upper and lower risk level monotonic trend component graphs are shown with the MK trend straight-line based on the classical Sen slope concept. It is observed that the median level CETA trends are in good agreement with the MK trends and although there are some numerical slope discrepancies especially at Izmir, Şanlıurfa and Samsun meteorology stations, but they have all decreasing general tendencies. As mentioned above, these three stations could not pass the serial independence assumption requirement of MK and therefore CETA method results can be regarded as better representations.

## 5. Discussion

After all what have been presented in the previous sections, one can flourish the following privileged points for the CETA method.

1. CETA is an empirical crossing procedure in search of a set of trends within a given hydro-meteor-climatic series without any restrictive assumption,
2. CETA helps to define possible monotonic trend component for any point within the variation domain of available data, which has been exemplified in this paper at three levels, namely median (50%), lower and upper (10%) risk levels. Hence, it is possible to identify monotonic trend lines at a set of points contrary to the classical approaches, where there is only one possibility,
3. It is noticed that each trend has different intercept and especially slope values, whereas in the classical approaches, say in the MK trend test, there is only a single trend calculated according to Sen (1968) slope procedure,
4. On the average levels (arithmetic mean or median) the CETA and MK trend calculations appear rather close to each other with practically acceptable different slopes. It seems that the CETA results have more confidence, because there is no restrictive assumption as for the underlying PDF and serial independence,
5. Both low and high extreme values have different trend lines with their own intercepts and slopes,
6. As for the drought duration period calculations CETA method provides more reliable calculation bases along each trend line,

## 6. Conclusions

Monotonic trend identification in any given hydro-meteor-climatic records has gained importance since almost three decades and it seems that this tendency will continue in the future due to the global warming and its consequent result of climate change and variability impacts. This paper presents a new approach for trend detection based on the crossing empirical trend

analysis (CETA) conception, which depends on the maximum crossing number along a straight-line that passes through a given pivot point. The trend slope is equal to the tangent of this straight line. This method does not require any restrictive assumption and provides trend slope calculation different than the classical approaches. The CETA method is a non-parametric approach that does not depend on the classical statistical parameters. The proposed method yields trend component at any given point that remains within the variation domain of the hydro-meteor-climatic record. In this paper, three pivot points are considered as the median corresponding to 50% risk level in addition to two extreme value risk levels each at the upper and lower 10% amounts of the fitted theoretical probability distribution function (PDF). The classical trend identification methodologies yield a single monotonic trend within a given time series with its single slope value, but CETA yields different slope values for any trend straight-line with the maximum crossing number. The application of the CETA methodology is presented for seasonal monthly and annual discharge records from Danube River in addition to seven meteorology station annual total precipitation amounts from seven geographical regions of Turkey. The comparison of the classical Mann-Kendall (MK) trend straight-line with the classical Sen's slope method calculation indicates that CETA yields different, but rather close trend lines to MK lines.

## Declarations

### Conflict of interest

Author has no conflict of interest.

### Funding Statement

This research has not received any funding from any source.

### Author's Contribution

Only the author contributed to this article.

### Availability of data and material

The data are obtained from internet for Danube River and for Turkish rainfall records from the General Directorate of Meteorology, Turkey

### Code availability

The code is written in Matlab by the authors and it is available on request.

### Ethics approval

This article abides by copyright laws.

### Consent to participate

Not applicable.

### Consent for publication

Not applicable.

## References

Ahmadi, F., Nazeri Tahroudi, M., Mirabbasi, R., Khalili, K., Jhajharia, D., (2018). [Spatiotemporal trend and abrupt change analysis of temperature in Iran](https://doi.org/10.1002/met.1694). *Meteorological Applications*, Vol. 25(2): 3124-321. <https://doi.org/10.1002/met.1694>

Bayazit, M., and Önöz, B., (2007). To pre-whiten or not to pre-whiten in trend analysis?, *Hydrological Sciences Journal*, Vol. 5(4),: 611-624, DOI: 10.1623/hysj.52.4.611.

Chakravarty, I.M. et al . (1967). *Handbook of Methods of Applied Statistics*, Willey Publications, New York.

Hamed, K. H., Rao, A. R., (1998). [A modified Mann-Kendall trend test for autocorrelated data](#). *Journal of Hydrology*, Vol. 204(1): 182-196. [https://doi.org/10.1016/S0022-1694\(97\)00125-X](https://doi.org/10.1016/S0022-1694(97)00125-X)

Hirsch, R. M., Alexander, R. B., Smith, R. A., (1991). [Selection of methods for the detection and estimation of trends in water quality](#). *Water Resources Research*, Vol. 27(5): 803-813. <https://doi.org/10.1029/91WR00259>

Khalili, K., Tahoudi, M. N., Mirabbasi, R., Ahmadi, F., (2016). [Investigation of spatial and temporal variability of precipitation in Iran over the last half century](#). *Stochastic Environmental Research and Risk Assessment*, Vol. 30(4): 1205-1221. <https://doi.org/10.1007/s00477-015-1095-4>

Kendall, M. G., (1975). *Rank Correlation Methods*, 4th ed., Charles Griffin: London.

Mann, H. B., (1945). Non-parametric test against trend, *Econometrica*, Vol. 13: 245-259.

Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., and Stouffer, R.J., (2008). Stationarity Is Dead: Whither Water Management? *Science*, Vol. 319: 573-574.

Sen, P. K., (1968). Estimates of the regression coefficient based on Kendall's Tau, *J. Am. Stat. Assoc.*, Vol. 63: 1379-1389.

Spearman, C., (1904). The proof and measurement of association between two things. *Amer. J. Psychol.* Vol. **15**: 72–101. doi:10.2307/1412159.

Şen, Z., (2012). [Innovative trend analysis methodology](#). *Journal of Hydrologic Engineering*, Vol. 17 (9); 1042-1046:

Şen, Z., (2017). Hydrological trend analysis with innovative and over-whitening procedures. *Hydrological Sciences Journal*, Vol. 62(2): 294-305, DOI: [10.1080/02626667.2016.1222533](https://doi.org/10.1080/02626667.2016.1222533).

Şen, Z., (2018). Crossing trend analysis methodology and application for Turkish rainfall records. *Theor Appl Climatol*, Vol. 131: 285–293. <https://doi.org/10.1007/s00704-016-1980-x>

Yue, S., Pilon, P., Cavadias, G., (2002). [Power of the Mann–Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series](#). *Journal of Hydrology*, Vol. 259(1): 254-271. [https://doi.org/10.1016/S0022-1694\(01\)00594-7](https://doi.org/10.1016/S0022-1694(01)00594-7)

## Figures

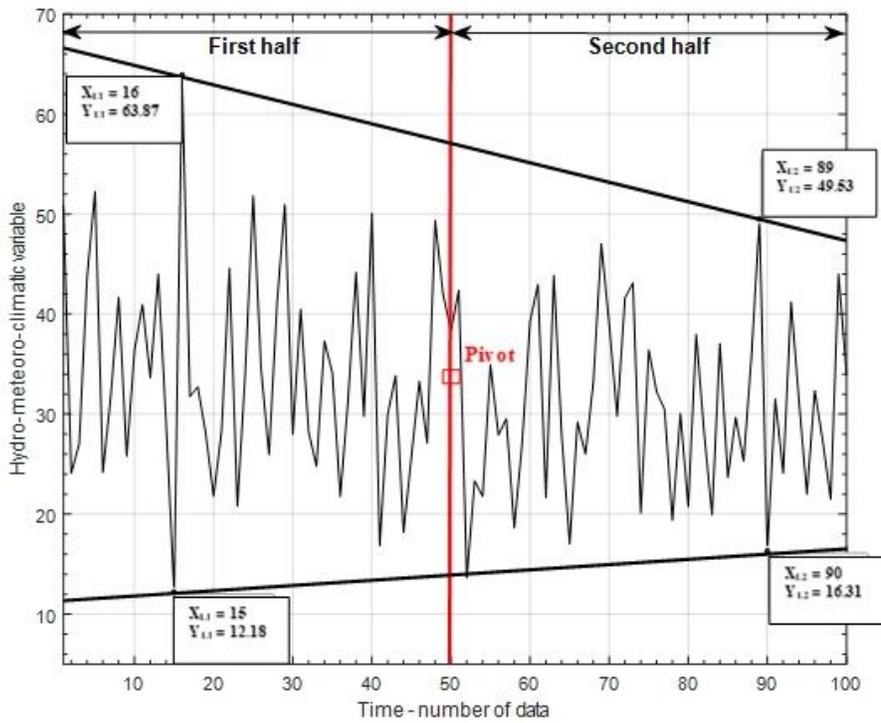


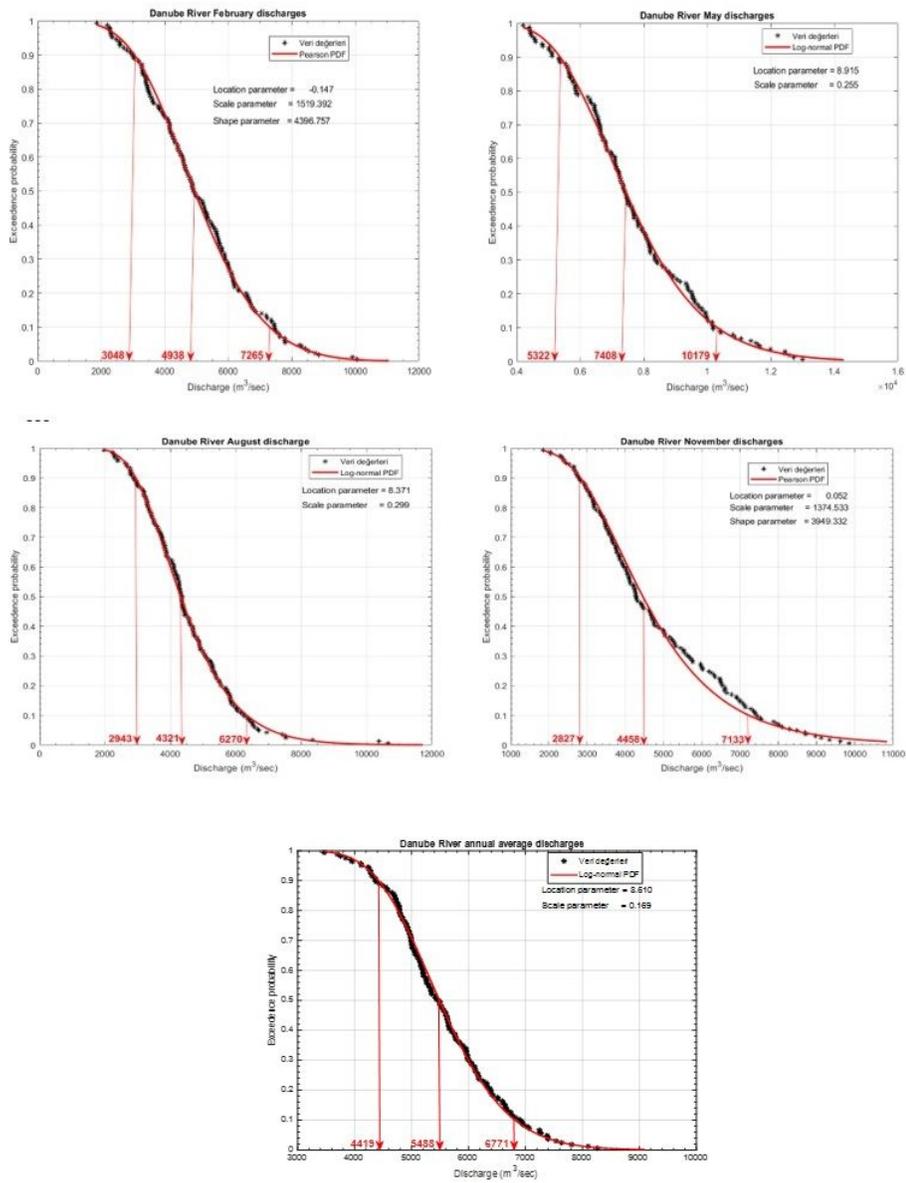
Figure 1

CETA method illustrative time series with relevant components



Figure 2

Seven geographic and climatological regions in Turkey. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.



**Figure 3**

Monthly PDF types, parameters and three risk levels

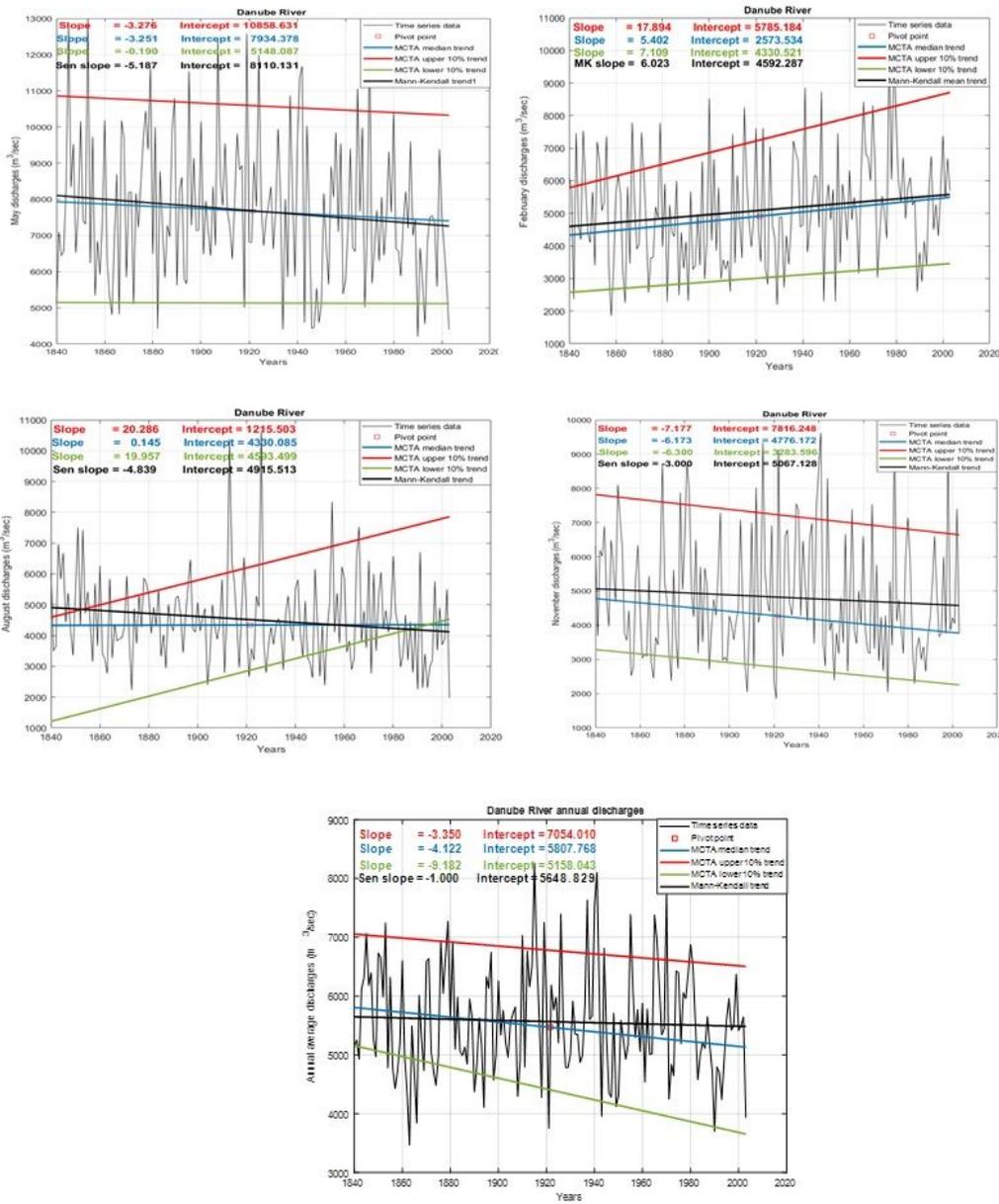


Figure 4

Four months and annual CETA trends

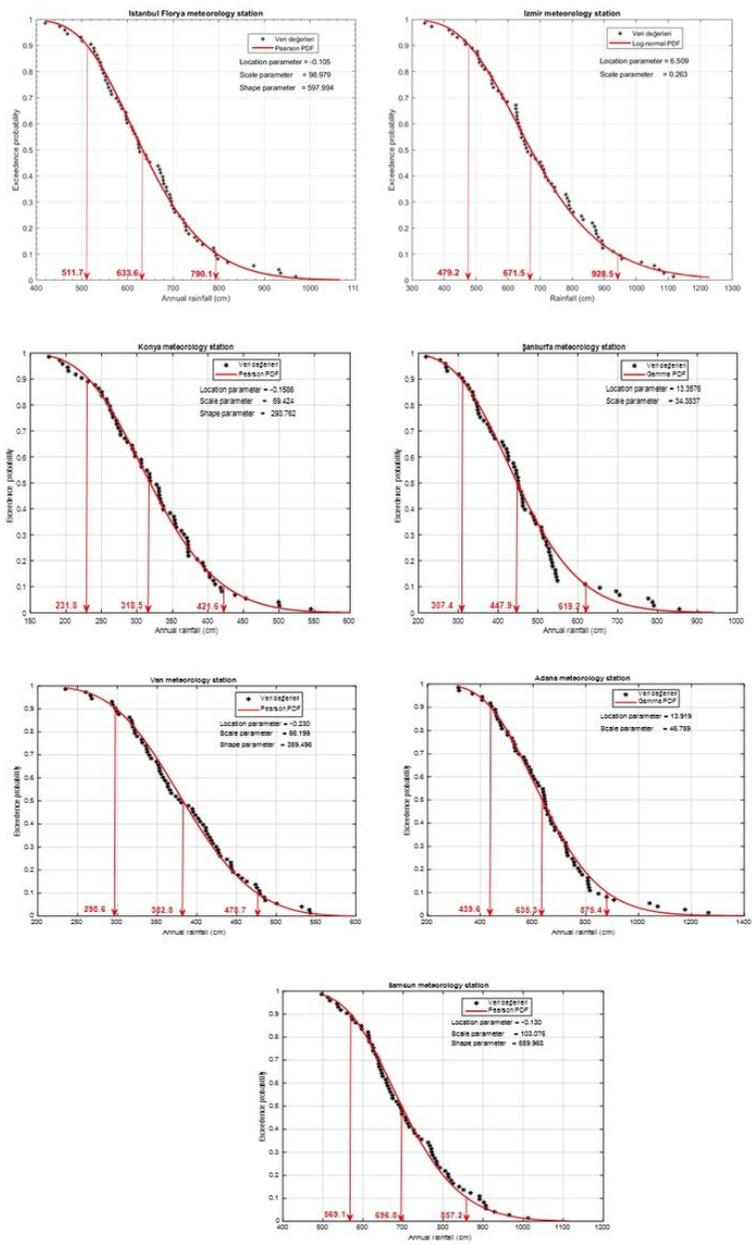


Figure 5

Annual precipitation PDF types, parameters and three risk levels

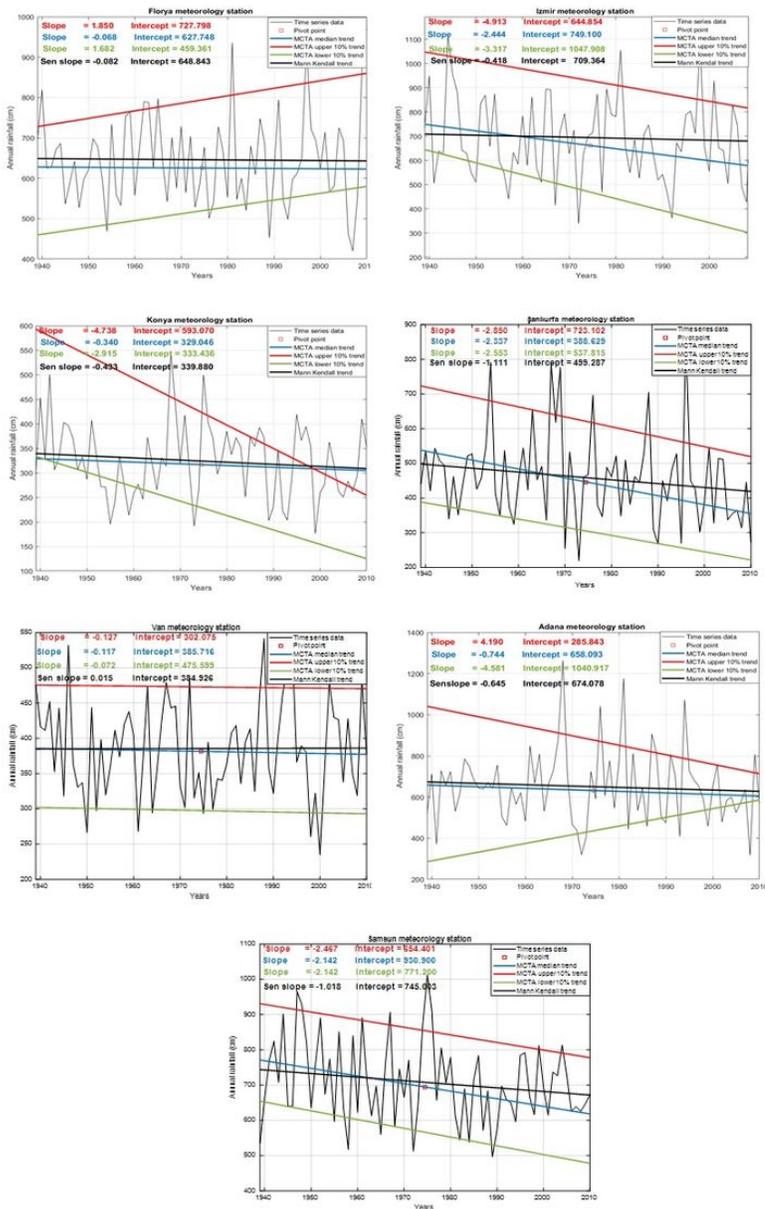


Figure 6

CETA annual precipitation trends at three risk levels